

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/709,585

Source: _____

Date Processed by STIC: _____

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 05/16/2005

PATENT APPLICATION: US/09/709,585

TIME: 11:32:15

Input Set : N:\Crf3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Vinik, Aaron I.

6 Pittenger, Gary L.

7 Rafaeloff, Ronit

8 Rosenberg, Lawrence

9 Duguid, William P.

11 (ii) TITLE OF INVENTION: INGAP PROTEIN INVOLVED IN PANCREATIC
12 ISLET NEOGENESIS

14 (iii) NUMBER OF SEQUENCES: 7

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: Banner & Allegretti

18 (B) STREET: 1001 G Street, N.W.

19 (C) CITY: Washington

20 (D) STATE: D.C.

21 (E) COUNTRY: US

22 (F) ZIP: 20001-4597

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Floppy disk

26 (B) COMPUTER: IBM PC compatible

27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30 (vi) CURRENT APPLICATION DATA:

C--> 31 (A) APPLICATION NUMBER: US/09/709,585

C--> 32 (B) FILING DATE: 13-Nov-2000

W--> 38 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: US/08/401,530

37 (B) FILING DATE: 24-OCT-1996

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Kagan, Sarah A.

42 (B) REGISTRATION NUMBER: 32,141

43 (C) REFERENCE/DOCKET NUMBER: 00570.48743

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: 202-508-9100

47 (B) TELEFAX: 202-508-9299

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 747 base pairs

54 (B) TYPE: nucleic acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: cDNA

RAW SEQUENCE LISTING

DATE: 05/16/2005

PATENT APPLICATION: US/09/709,585

TIME: 11:32:15

Input Set : N:\Crif3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

```

60      (iii) HYPOTHETICAL: NO
62      (iv) ANTI-SENSE: NO
64      (vi) ORIGINAL SOURCE:
65          (A) ORGANISM: Cricetulus
67      (ix) FEATURE:
68          (A) NAME/KEY: CDS
69          (B) LOCATION: 20..541
72      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
74 CTGCAAGACA GGTACCATG ATG CTT CCC ATG ACC CTC TGT AGG ATG TCT TGG      52
75          Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp
76          1          5          10
78 ATG CTG CTT TCC TGC CTG ATG TTC CTT TCT TGG GTG GAA GGT GAA GAA      100
79 Met Leu Leu Ser Cys Leu Met Phe Leu Ser Trp Val Glu Gly Glu Glu
80          15          20          25
82 TCT CAA AAG AAA CTG CCT TCT TCA CGT ATA ACC TGT CCT CAA GGC TCT      148
83 Ser Gln Lys Lys Leu Pro Ser Ser Arg Ile Thr Cys Pro Gln Gly Ser
84          30          35          40
86 GTA GCC TAT GGG TCC TAT TGC TAT TCA CTG ATT TTG ATA CCA CAG ACC      196
87 Val Ala Tyr Gly Ser Tyr Cys Tyr Ser Leu Ile Leu Ile Pro Gln Thr
88          45          50          55
90 TGG TCT AAT GCA GAA CTA TCC TGC CAG ATG CAT TTC TCA GGA CAC CTG      244
91 Trp Ser Asn Ala Glu Leu Ser Cys Gln Met His Phe Ser Gly His Leu
92 60          65          70          75
94 GCA TTT CTT CTC AGT ACT GGT GAA ATT ACC TTC GTG TCC TCC CTT GTG      292
95 Ala Phe Leu Leu Ser Thr Gly Glu Ile Thr Phe Val Ser Ser Leu Val
96          80          85          90
98 AAG AAC AGT TTG ACG GCC TAC CAG TAC ATC TGG ATT GGA CTC CAT GAT      340
99 Lys Asn Ser Leu Thr Ala Tyr Gln Tyr Ile Trp Ile Gly Leu His Asp
100          95          100          105
102 CCC TCA CAT GGT ACA CTA CCC AAC GGA AGT GGA TGG AAG TGG AGC AGT      388
103 Pro Ser His Gly Thr Leu Pro Asn Gly Ser Gly Trp Lys Trp Ser Ser
104          110          115          120
106 TCC AAT GTG CTG ACC TTC TAT AAC TGG GAG AGG AAC CCC TCT ATT GCT      436
107 Ser Asn Val Leu Thr Phe Tyr Asn Trp Glu Arg Asn Pro Ser Ile Ala
108          125          130          135
110 GCT GAC CGT GGT TAT TGT GCA GTT TTG TCT CAG AAA TCA GGT TTT CAG      484
111 Ala Asp Arg Gly Tyr Cys Ala Val Leu Ser Gln Lys Ser Gly Phe Gln
112 140          145          150          155
114 AAG TGG AGA GAT TTT AAT TGT GAA AAT GAG CTT CCC TAT ATC TGC AAA      532
115 Lys Trp Arg Asp Phe Asn Cys Glu Asn Glu Leu Pro Tyr Ile Cys Lys
116          160          165          170
118 TTC AAG GTC TAGGGCAGTT CTAATTTCAA CAGAGAGCAA GCTCTGCCTA CACACCCACA      591
119 Phe Lys Val
122 CCAATTCCCT TATATCATCT CTGCTGTTTT TCCTTGAAAT TATTATGAAG CTCACATGGA      651
125 CAAGGAAGCA AGTATGAGGA TTCACTCAGG ATATCAGTAT ATTCTGTGGT GGCTGTAACC      711
128 TAAAGGCTCA GAGAACAAAA ATAAAATGTC ATCAAC      747
131 (2) INFORMATION FOR SEQ ID NO: 2:
133      (i) SEQUENCE CHARACTERISTICS:
134          (A) LENGTH: 174 amino acids

```

RAW SEQUENCE LISTING

DATE: 05/16/2005

PATENT APPLICATION: US/09/709,585

TIME: 11:32:15

Input Set : N:\Crif3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

```

135          (B) TYPE: amino acid
136          (D) TOPOLOGY: linear
138          (ii) MOLECULE TYPE: protein
140          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
142 Met  Leu  Pro  Met  Thr  Leu  Cys  Arg  Met  Ser  Trp  Met  Leu  Leu  Ser  Cys
143   1      5      10      15
145 Leu  Met  Phe  Leu  Ser  Trp  Val  Glu  Gly  Glu  Ser  Gln  Lys  Lys  Leu
146      20      25      30
148 Pro  Ser  Ser  Arg  Ile  Thr  Cys  Pro  Gln  Gly  Ser  Val  Ala  Tyr  Gly  Ser
149      35      40      45
151 Tyr  Cys  Tyr  Ser  Leu  Ile  Leu  Ile  Pro  Gln  Thr  Trp  Ser  Asn  Ala  Glu
152      50      55      60
154 Leu  Ser  Cys  Gln  Met  His  Phe  Ser  Gly  His  Leu  Ala  Phe  Leu  Leu  Ser
155      65      70      75      80
157 Thr  Gly  Glu  Ile  Thr  Phe  Val  Ser  Ser  Leu  Val  Lys  Asn  Ser  Leu  Thr
158      85      90      95
160 Ala  Tyr  Gln  Tyr  Ile  Trp  Ile  Gly  Leu  His  Asp  Pro  Ser  His  Gly  Thr
161      100     105     110
163 Leu  Pro  Asn  Gly  Ser  Gly  Trp  Lys  Trp  Ser  Ser  Ser  Asn  Val  Leu  Thr
164      115     120     125
166 Phe  Tyr  Asn  Trp  Glu  Arg  Asn  Pro  Ser  Ile  Ala  Ala  Asp  Arg  Gly  Tyr
167      130     135     140
169 Cys  Ala  Val  Leu  Ser  Gln  Lys  Ser  Gly  Phe  Gln  Lys  Trp  Arg  Asp  Phe
170      145     150     155     160
172 Asn  Cys  Glu  Asn  Glu  Leu  Pro  Tyr  Ile  Cys  Lys  Phe  Lys  Val
173      165     170
175 (2) INFORMATION FOR SEQ ID NO: 3:
177   (i) SEQUENCE CHARACTERISTICS:
178       (A) LENGTH: 175 amino acids
179       (B) TYPE: amino acid
180       (D) TOPOLOGY: linear
182   (ii) MOLECULE TYPE: protein
184   (iv) ANTI-SENSE: NO
186   (vi) ORIGINAL SOURCE:
187       (A) ORGANISM: Rattus rattus
190   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
192 Met  Leu  His  Arg  Leu  Ala  Phe  Pro  Val  Met  Ser  Trp  Met  Leu  Leu  Ser
193   1      5      10      15
195 Cys  Leu  Met  Leu  Leu  Ser  Gln  Val  Gln  Gly  Glu  Asp  Ser  Pro  Lys  Lys
196      20      25      30
198 Ile  Pro  Ser  Ala  Arg  Ile  Ser  Cys  Pro  Lys  Gly  Ser  Gln  Ala  Tyr  Gly
199      35      40      45
201 Ser  Tyr  Cys  Tyr  Ala  Leu  Phe  Gln  Ile  Pro  Gln  Thr  Trp  Phe  Asp  Ala
202      50      55      60
204 Glu  Leu  Ala  Cys  Gln  Lys  Arg  Pro  Glu  Gly  His  Leu  Val  Ser  Val  Leu
205      65      70      75      80
207 Asn  Val  Ala  Glu  Ala  Ser  Phe  Leu  Ala  Ser  Met  Val  Lys  Asn  Thr  Gly
208      85      90      95
210 Asn  Ser  Tyr  Gln  Tyr  Ile  Trp  Ile  Gly  Leu  His  Asp  Pro  Thr  Leu  Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/709,585

DATE: 05/16/2005

TIME: 11:32:15

Input Set : N:\Crf3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

```

211          100          105          110
213 Gly Glu Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Asn Asp Ile Met
214          115          120          125
216 Asn Tyr Val Asn Trp Glu Arg Asn Pro Ser Thr Ala Leu Asp Arg Gly
217          130          135          140
219 Phe Cys Gly Ser Leu Ser Arg Ser Ser Gly Phe Leu Arg Trp Arg Asp
220          145          150          155          160
222 Thr Thr Cys Glu Val Lys Leu Pro Tyr Val Cys Lys Phe Thr Gly
223          165          170          175
225 (2) INFORMATION FOR SEQ ID NO: 4:
227 (i) SEQUENCE CHARACTERISTICS:
228 (A) LENGTH: 175 amino acids
229 (B) TYPE: amino acid
230 (D) TOPOLOGY: linear
232 (ii) MOLECULE TYPE: protein
234 (vi) ORIGINAL SOURCE:
235 (A) ORGANISM: Homo sapiens
238 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
240 Met Leu Pro Pro Met Ala Leu Pro Ser Val Ser Trp Met Leu Leu Ser
241 1 5 10 15
243 Cys Leu Met Leu Leu Ser Gln Val Gln Gly Glu Glu Pro Gln Arg Glu
244 20 25 30
246 Leu Pro Ser Ala Arg Ile Arg Cys Pro Lys Gly Ser Lys Ala Tyr Gly
247 35 40 45
249 Ser His Cys Tyr Ala Leu Phe Leu Ser Pro Lys Ser Trp Thr Asp Ala
250 50 55 60
252 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly Asn Leu Val Ser Val Leu
253 65 70 75 80
255 Ser Gly Ala Glu Gly Ser Phe Val Ser Ser Leu Val Lys Ser Ile Gly
256 85 90 95
258 Asn Ser Tyr Ser Tyr Val Trp Ile Gly Leu His Asp Pro Thr Gln Gly
259 100 105 110
261 Thr Glu Pro Asn Gly Glu Gly Trp Glu Trp Ser Ser Ser Asp Val Met
262 115 120 125
264 Asn Tyr Phe Ala Trp Glu Arg Asn Pro Ser Thr Ile Ser Ser Pro Gly
265 130 135 140
267 His Cys Ala Ser Leu Ser Arg Ser Thr Ala Phe Leu Arg Trp Lys Asp
268 145 150 155 160
270 Tyr Asn Cys Asn Val Arg Leu Pro Tyr Val Cys Lys Phe Thr Asp
271 165 170 175
273 (2) INFORMATION FOR SEQ ID NO: 5:
275 (i) SEQUENCE CHARACTERISTICS:
276 (A) LENGTH: 174 amino acids
277 (B) TYPE: amino acid
278 (D) TOPOLOGY: linear
280 (ii) MOLECULE TYPE: protein
282 (vi) ORIGINAL SOURCE:
283 (A) ORGANISM: Rattus rattus
286 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/709,585

DATE: 05/16/2005

TIME: 11:32:15

Input Set : N:\Crif3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

```

288 Met Leu Pro Arg Val Ala Leu Thr Thr Met Ser Trp Met Leu Leu Ser
289 1 5 10 15
291 Ser Leu Met Leu Leu Ser Gln Val Gln Gly Glu Asp Ala Lys Glu Asp
292 20 25 30
294 Val Pro Thr Ser Arg Ile Ser Cys Pro Lys Gly Ser Arg Ala Tyr Gly
295 35 40 45
297 Ser Tyr Cys Tyr Ala Leu Phe Ser Val Ser Lys Ser Trp Phe Asp Ala
298 50 55 60
300 Asp Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Val Leu
301 65 70 75 80
303 Ser Gly Ser Glu Ala Ser Phe Val Ser Ser Leu Ile Lys Ser Ser Gly
304 85 90 95
306 Asn Ser Gly Gln Asn Val Trp Ile Gly Leu His Asp Pro Thr Leu Gly
307 100 105 110
309 Gln Glu Pro Asn Arg Gly Gly Trp Glu Trp Ser Asn Ala Asp Val Met
310 115 120 125
312 Asn Tyr Phe Asn Trp Glu Thr Asn Pro Ser Ser Val Ser Gly Ser His
313 130 135 140
315 Cys Gly Thr Leu Thr Arg Ala Ser Gly Phe Leu Arg Trp Arg Glu Asn
316 145 150 155 160
318 Asn Cys Ile Ser Glu Leu Pro Tyr Val Cys Lys Phe Lys Ala
319 165 170

```

321 (2) INFORMATION FOR SEQ ID NO: 6:

322 (i) SEQUENCE CHARACTERISTICS:

323 (A) LENGTH: 174 amino acids

324 (B) TYPE: amino acid

325 (D) TOPOLOGY: linear

326 (ii) MOLECULE TYPE: protein

327 (vi) ORIGINAL SOURCE:

328 (A) ORGANISM: Rattus rattus

329 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

336 Met Leu Pro Arg Leu Ser Phe Asn Asn Val Ser Trp Thr Leu Leu Tyr
337 1 5 10 15
339 Tyr Leu Phe Ile Phe Gln Val Arg Gly Glu Asp Ser Gln Lys Ala Val
340 20 25 30
342 Pro Ser Thr Arg Thr Ser Cys Pro Met Gly Ser Lys Ala Tyr Arg Ser
343 35 40 45
345 Tyr Cys Tyr Thr Leu Val Thr Leu Lys Ser Trp Phe Gln Ala Asp
346 50 55 60
348 Leu Ala Cys Gln Lys Arg Pro Ser Gly His Leu Val Ser Ile Leu Ser
349 65 70 75 80
351 Gly Gly Glu Ala Ser Phe Val Ser Ser Leu Val Thr Gly Arg Val Asn
352 85 90 95
354 Asn Asn Gln Asp Ile Trp Ile Trp Leu His Asp Pro Thr Met Gly Gln
355 100 105 110
357 Gln Pro Asn Gly Gly Gly Trp Glu Trp Ser Asn Ser Asp Val Leu Asn
358 115 120 125
360 Tyr Leu Asn Trp Asp Gly Asp Pro Ser Ser Thr Val Asn Arg Gly Asn
361 130 135 140

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/709,585

DATE: 05/16/2005

TIME: 11:32:16

Input Set : N:\Crf3\RULE60\09709585.raw.txt

Output Set: N:\CRF4\05162005\I709585.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
 L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
 L:38 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)